RECOGNITION OF CROPS AND SOILS BY SPOT DENSITY MEASUREMENTS OF IMAGERY

by

Gerald D. Nelson
Remote Sensing Institute
South Dakota State University
Brookings, South Dakota

INTRODUCTION

The goals of the pattern recognition work in progress at the Remote Sensing Institute at South Dakota State University are:

1. To develop methods that are useful to the analyses, feature definition, feature selection, and classification of remotely sensed data, and
2. to determine the usefulness of spot density measurements of Ektachrome infrared film for use as features to classify crops from altitudes of 60,000 feet (NASA flight), and crops and soils from 14,000 feet (Remote Sensing Institute flight).

To develop automatic methods to aid as well as to classify the data resulting from imagery requires an extensive effort be supplied to feasibility studies. The pattern recognition feasibility study areas include:

Imagery measurements
Data compression methods
Feature definition
Feature extraction
Feature analysis
Feature selection
Classification
Encoding of classification results
Color display techniques

In order to develop a satisfactory pattern recognition system requires careful interfacing of all the feasibility study areas. Naturally, one recognizes this not as a new concept, but as a system approach to design which has significant merit.

The second goal is to determine the usefulness of spot density measurements of the imagery for different films and flight altitudes. The tasks encompassed have been all but the data compression, encoding of classification results and color display techniques. The feasibility study area results are discussed in this paper.

A major value of automatic or even semi-automatic pattern recognition techniques lies in the area of making repetitive measurements, numerical calculations and decisions without tiring as does the human. The trained human at the present time is still better qualified as a decision maker than any machine which uses density measurements as the features. Therefore, the research done in the area of pattern recognition for special tasks is still a search for reliable measurements which will provide adequate classification results to make the use of pattern recognition techniques economically feasible. It is also desirable that the results be as accurate as those of a good human photo interpreter. Implied, as a goal, in pattern recognition research is that the methods be computationally efficient.

STUDY AREA

The description of the study area is presented by a set of three photos and their associated transparent overlays which outline the soils and/or crops. This set of three photos does not include all of the fields on which densitometer measurements were made.

The soil study at the present time has been restricted to two soils denoted as soil A and B. The two soils studied are outlined on the Ektachrome infrared photo contact printed from the transparency. This photo was taken at an altitude of 14,000 feet.

A study of the recognition of crops from altitudes of 60,000 feet and 14,000 feet were also conducted. The crop identification of fields from 60,000 feet is presented as Figure 2. Ektachrome infrared film was used.

The study of crops from an altitude of 14,000 feet is easier because the fields are larger and easier for the human to recognize. Also more measurements can be made per field. The identification of crops from 14,000 feet is presented as Figure 3.

EXPERIMENTAL METHODS AND PROCEDURES

This section includes discussions of ground truth, densitometer measurements, pattern recognition computer programs, and a proposed hardware pattern classifier.

GROUND TRUTH

A "ground-truth" mission was conducted on July 4, 1969. An identification of 533 fields in three separate flight coverage areas was made. For the pattern recognition studies, nineteen classes were assigned and are presented in Table I. This ground truth data has been mapped, coded, and field numbered to enable coordination of derived data from more than one source.

DENSITOMETER MEASUREMENTS

The feasibility studies reported on within this report are based on the information contained in the measurements made with a densitometer with four different filters. The Macbeth densitometer was used with a one millimeter spot size. On the 14,000 foot imagery approximately 20 spot density measurements with each filter were made within a field, whereas only five spot density measurements were made with each filter within a field on the 60,000 foot imagery.

Another instrument, the Spatial Data system, can also be used to make density measurements. This is an instrument which uses a vidicon to sense the light transmitted through the film. The result is color encoded into as many as 32 colors. An advantage of this system is the speed at which the data is encoded.

PATTERN RECOGNITION COMPUTER PROGRAMS

There exists a need to observe the structure of the density measurements which are the features. This can be done by generating a sample probability density function for each set of density measurements per crop or field. From these plots one can estimate the value of the feature for pairwise class or crop classification, but when trying to interpret this data for a many-class problem one has difficulty.

When no single feature appears adequate, then it is desirable to use pairs of features to discriminate among the classes. A scatter-plot is useful to estimate the separability of classes by pairs of features. However, in the scatter plot the frequency of occurrence of each point is not presented, but can be determined by the list which is called the overprint record. The significant factor to be determined is the amount of overlap.

Another computer program which is helpful and should be used before a classification study is made determines the number of modes present in the data on a per class basis. Actually, the sample probability density function and scatter plots provide information as to the number of modes. However, the output of this program is more detailed than either of the other two. The output consists of the radius of each mode, the pairwise distance between crops or soils for each feature, and the total distance between crops for each of the features.

The classifier implemented as a computer program is based on determination of a matrix B which provides a least-squares mapping of the class vector estimate toward a set of orthonormal class vectors [1]. The minimization problem which is solved determines B by the minimization of the squared distance between the class orthonormal vectors and the class vector estimates.

An event or sample represented by the class vector estimate is assigned to that class whose class vector is closest, in a Euclidian sense, to the mapped feature vector which is

$$d = B\,x_A$$

where x_A is the n+1 dimensional augmented feature vector

$$x_A = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ xn-1 \\ xn \\ \hline -1 \end{bmatrix}$$

The decision rule is to select class i

if $d_i > d_j$ for all $j \neq i$.

The decision vector is also represented by the equation

$$d = [P_i]\ [\overline{x}^i - \overline{x}]_R^T \phi^{-1} [x - \overline{x}] + [P_i]$$

$$\text{where } [P_i] = \begin{bmatrix} p_1 & o & o \\ o & p_2 & \\ o & & p_k \end{bmatrix}$$

are the a priori probabilities of each class occurring or a set of weights since usually the a priori probabilities are unknown.

$\overline{x}^i$ is the mean vector of the i^{th} class,
$\overline{x}$ is the mean vector of all classes,
ϕ is the sample covariance matrix and is calculated according to

$$\phi = [\overline{xx^T} - \overline{x}\overline{x}^T$$

ϕ^{-1} is the inverse of the sample covariance matrix.

The normal process to determine the classifier structure is to supply a "training" set of feature vectors. From this "training" set $\overline{x}^i$, $\overline{x}$, $[P_i]$, ϕ, and ϕ^{-1} are calculated. The only unknown term remaining in the equation for the decision vector is x, the feature vector. Therefore, at this point, the classifier is trained and either the training set or "testing" set of data is supplied to the classifier program. The decision vector is calculated for each feature vector and the classification result determined by selecting the subscript of the largest element of the decision vector as the correct class number. The result of this process is the confusion matrix which represents the score attained in the classification process. As an example consider Table II.

Ninety-five percent of the class three feature vectors were classified as class three, one percent were classified as class one, and four percent as class two.

SPECIAL CLASSIFICATION HARDWARE

The decision vector can also be written as in the next equation,

$$d = [P_i] \; [(\overline{x}^i-\overline{x})_R \phi^{-1} x-(\overline{x}^i-\overline{x})_R \phi^{-1}\overline{x}+1]$$

$$d = Ax-A\overline{x}+1$$

where $A = [P_i] \; [\overline{x}^i-\overline{x}]\phi^{-1}$

$$d = A \; (x-\overline{x}) + [P_i]$$

From the last equation it is obvious that to build a special purpose classifier once the training phase is completed is relatively easy. To build the hardware required to train the classifier is not so easy. A storage is required for the matrix $[P_i]$ and the vector $\overline{x}$. The difference between the feature vector x and the average feature vector of all classes $\overline{x}$ is formed and the result multiplied by the matrix A. To this product are added the weights $[P_i]$ and then the largest element of the decision vector is determined and the decision recorded or announced by indicator lights. A block diagram of the classifier is shown in Figure 4.

This classifier is proposed as a slow speed system which could effectively demonstrate the decision at boundaries, or other selected spots on a film once the classifier was trained. The main advantages are that the density measurements do not have to be recorded, keypunched, verified, or positional information encoded so that the measurement spot can be located after the computer classification results are printed. The classification results could be determined as rapidly as the human can make them, and the human stores the positional information. In fact, the human acts as an adaptive sampler and determines results only at the location of special interest to him.

PRELIMINARY SOIL IDENTIFICATION EXPERIMENT

Two soils, referred to as soil A and B, were identified by Dr. Frazee on Ektachrome infrared film exposed at an altitude of 14,000 feet. To determine if density measurements could be used as features to recognize these soils the following tasks of the experiment were conducted.

1. Measurement of 160 spot density readings per soil type with a one millimeter aperture on the Macbeth densitometer. Each of the four filters were used, neutral or visible, red, green and blue.
2. Plot of the sample probability density function for all filters and each soil.
3. Plot of all two-dimensional scatter plots for the two soils.
4. Classification into two classes based on the four density readings per location.

The plot of the sample probability density function for each feature in Figure 5 indicates the best individual feature to discriminate between these soils is the red filter density measurement.

To determine the best pair of features to discriminate between these soils it is necessary to consider the scatter plots shown in Figure 6. There are several pairs of features that appear they could be used effectively. They are the blue-green and blue-red filters. These plots indicate that there is no overlap of the density measurements since there are no + signs indicated.

To classify the samples of a program K-class I was used and all four features were used. The confusion matrix or score matrix is presented as Table III. This result indicates that in the four-dimensional space the two soils are almost linearly separable.

The use of the Spatial Data system for quantizing or level slicing should be used with a red filter to get the best results for these soils with one filter. Figure 7 illustrates that a neutral filter on the vidicon does not separate the two soils, but Figure 8 indicates that a red filter does separate the two soils.

One of the major problems in pattern recognition work is to determine the procedure for the selection of features. In the present case since there are only four features, an exhaustive search for the best solution is feasible. The classification results for each of the fifteen combinations features the rank ordering of the classification results are presented in Table IV.

One should note that the two worst features separately, blue provides 53.75 percent and green 82.19 percent, if used together as a pair they provide a correct classification of 98.75 percent.

The classification results as a function of the number of features used, are presented in Figure 9.

PRELIMINARY CROP IDENTIFICATION EXPERIMENTS

This preliminary study consisted of using Ektachrome infrared film exposed from 14,000 and 60,000 feet. As was the case with the soil study, spot density measurements were used. A one millimeter aperture on the MacBeth densitometer was used with four filters which included neutral or visible, red, green and blue.

60,000 FEET, EKTACHROME INFRARED FILM

The classification results for the Ektachrome infrared film exposed at 60,000 feet are presented in Table V. The K-class I program was used to classify the crops. The low percentage of correct classification is believed to be due primarily to the large spot size used for imagery taken at 60,000 feet.

14,000 FEET, EKTACHROME INFRARED FILM

The Ektachrome infrared film exposed at 14,000 feet appears to be more useful to study than Ektachrome infrared at 60,000 feet. Corn, fallow, harvested wheat, and pasture grass were classified 69.5 percent correct as shown in Table VI.

The classifiers based on the use of alfalfa, wheat, harvested oats, and harvested alfalfa; and sorghum, oats, and hayland do not yield results which are as good as the other 14,000 feet Ektachrome infrared imagery, as shown in Tables VII and VIII. The percent of correct recognition is 20 and 62, respectively.

A classifier for six classes which are corn, fallow, harvested wheat, roadways, trees, and water are determined. The confusion matrix is presented as Table IX. The correct recognition rate for all classes is 75.5 percent. However, the fallow class is difficult to recognize. This difficulty could possibly be traced back to the ground truth definition of fallow which includes plowed and unplowed fields as well as fields with weeds. The poor classification results should be investigated by examination and comparison of the fields erroneously classified with those correctly classified.

The sample probability density functions for this six class problem are presented in Figure 10. I am sure that the human observer has a difficult time specifying the decision boundaries in this multi-class problem whereas in any two-class problem it may be quite easy. The classifier recognized 75 percent of these spot density measurements.

SUMMARY AND CONCLUSIONS

Computerized techniques and methods have been developed which were used to conduct preliminary soil and crop identification experiments. They will also be used to continue the study of classification and/or identification methods. However, additional methods which are assured to provide better results than reported in this report are also being developed [2].

The soil identification experiment was conducted by making densitometer measurements on Ektachrome infrared film exposed at 14,000 feet. The density measurements were analyzed by plotting sample probability density functions, two-dimensional scatter plots, and the use of K-class I to determine the complete set of classification results for one, two, three and four features.

Due to the presence of nineteen classes, crop identification experiments were more difficult to formulate. This is partially due to the computer core size which limits the number of classes, features and/or samples. However, the classes of corn, fallow, harvested wheat, roadways, trees and water were classified 75 percent correct as reported in Table IX.

The amount of data used to make a decision has a definite effect on the quality of the decision. To use spot density readings of the film is probably the most elementary or basic measurement to be used to determine the decision. However, some of the results are encouraging even though one anticipates better classification results if more data is used.

One of the significant problems associated with classifiers is that they are sensitive to the subset of classes used as well as the subset of features.

REFERENCES

[1] N. Zagalsky, "A New Formulation of a Classification Procedure," M. S. Thesis, University of Minnesota, March 1968.

[2] W. G. Wee, "On Feature Selection in a Class of Distribution - Free Pattern Classifiers," <u>IEEE Transactions on Information Theory</u>, Vol. IT-16, No. 1, pp 47-55, January 1970.

TABLE I.- THE CLASSIFICATIONS ESTABLISHED FOR PATTERN RECOGNITION

Code	Identification
1	Corn
2	Wheat
3	Oats
4	Alfalfa
5	Fallow
6	Sorghum
7	Pasture-grass
8	Barley
9	Harvested wheat
10	Harvested oats
11	Harvested alfalfa
12	Harvested barley
13	Slough
14	Brome
15	Hayland
16	Unknown
17	Roadways
18	Trees
19	Water

TABLE II.- SAMPLE CONFUSION MATRIX

		Number of Measurements	Percent		Classified As 1	2	3
Known As	1	100	99		99	1	0
	2	100	98		0	98	2
	3	100	95		1	4	95
Totals		300	97.33	Weights	.333	.333	.333

TABLE III.- CONFUSION MATRIX, 14,000 EKTACHROME INFRARED

		Number of Measurements	Percent		Classified As Soil A	B
Known As	A	160	99		99	1
	B	160	100		0	100
Totals		320	99	Weights	.500	.500

TABLE IV.- RANK ORDERING OF FEATURES

Individual Features

Red	99.375
Neutral	83.437
Green	82.187
Blue	53.750

Two Features

Neutral, Red	99.375
Neutral, Green	99.062
Red, Green	99.062
Green, Blue	98.750
Red, Blue	96.875
Neutral, Blue	94.062

Three Features

Neutral, Red, Blue	99.687
Red, Green, Blue	99.375
Neutral, Red, Green	99.375
Neutral, Green, Blue	98.750

Four Features

Neutral, Red, Green, Blue	99.687

TABLE V.- CONFUSION MATRIX - 60,000 FEET
Ektachrome Infrared Film

	Number of Measurements	Percent	Sorghum	Fallow	H. Oats	H. Alfalfa
Sorghum	65	32.3	21	29	11	4
Fallow	120	22.5	65	27	17	11
Harvested Oats	85	37.7	28	12	32	13
Harvested Alfalfa	50	24.0	15	16	7	12
			Weights			
Totals	320	28.7	.250	.250	.250	.250

TABLE VI.- CONFUSION MATRIX - 14,000 FEET
Ektachrome Infrared Film

	Number of Measurements	Percent	Corn	Fallow	H. Wheat	Pasture Grass
Corn	200	93.5	187	0	10	3
Fallow	200	79.0	0	158	42	0
H. Wheat	200	81.5	13	4	163	20
Pasture Grass	200	23.0	48	33	73	46
			Weights			
Totals	800	69.5	.250	.250	.250	.250

TABLE VII.- CONFUSION MATRIX - 14,000 FEET
Ektachrome Infrared Film

	Number of Measurements	Percent	Alfalfa	Wheat	H. Oats	H. Alfalfa
Alfalfa	100	36.0	36	0	61	3
Wheat	100	32.0	9	32	54	5
H. Oats	180	7.8	84	75	14	7
H. Alfalfa	160	16.3	44	52	38	26
			Weights			
Totals	540	20	.250	.250	.250	.250

TABLE VIII.- CONFUSION MATRIX - 14,000 FEET
Ektachrome Infrared Film

	Number of Measurements	Percent	Sorghum	Oats	Hayland
Sorghum	100	65.0	65	25	10
Oats	80	68.7	1	55	24
Hayland	60	40.0	18	18	24
			Weights		
Totals	240	62	.333	.333	.333

TABLE IX.- CONFUSION MATRIX - 14,000 FEET
Ektachrome Infrared Film

	Number of Measurements	Percent	Corn	Fallow	H. Wheat	Roadways	Trees	Water
Corn	200	73.0	146	0	1	6	47	0
Fallow	200	54.5	0	109	19	33	0	39
H. Wheat	200	72.5	17	1	145	37	0	0
Roadways	200	81.0	3	14	21	162	0	0
Trees	200	89.0	21	0	1	0	178	0
Water	200	83.0	3	29	1	1	0	166
			Weights					
Totals	1200	75.5	.167	.167	.167	.167	.167	.167

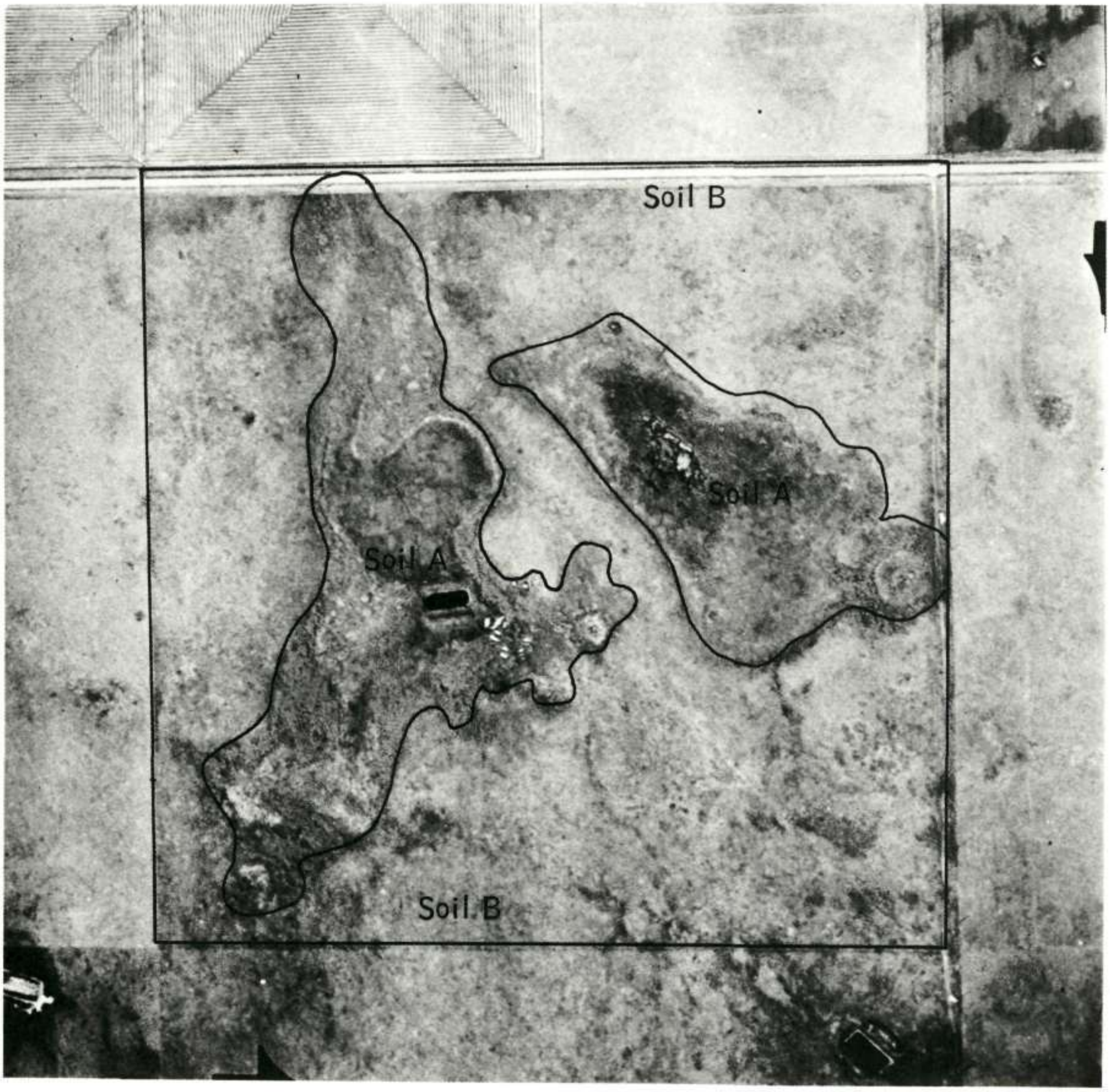


Figure 1. Soils A and B -- Pattern recognition study used data measured with Macbeth densitometer and Spatial Data system.

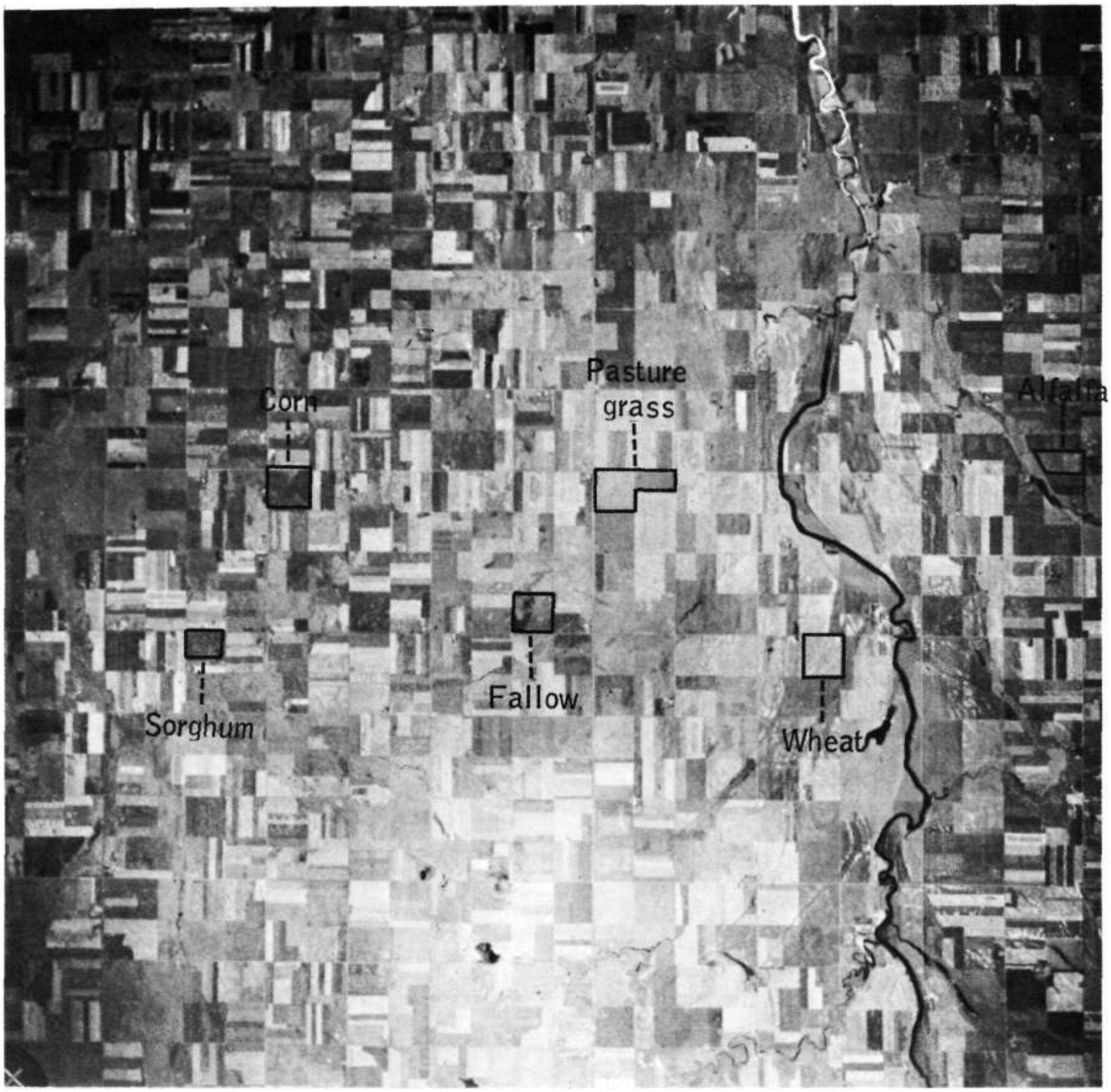


Figure 2. National Aeronautics and Space Administration imagery taken on August 8, 1969 over South Dakota at an altitude of 60,000 feet.

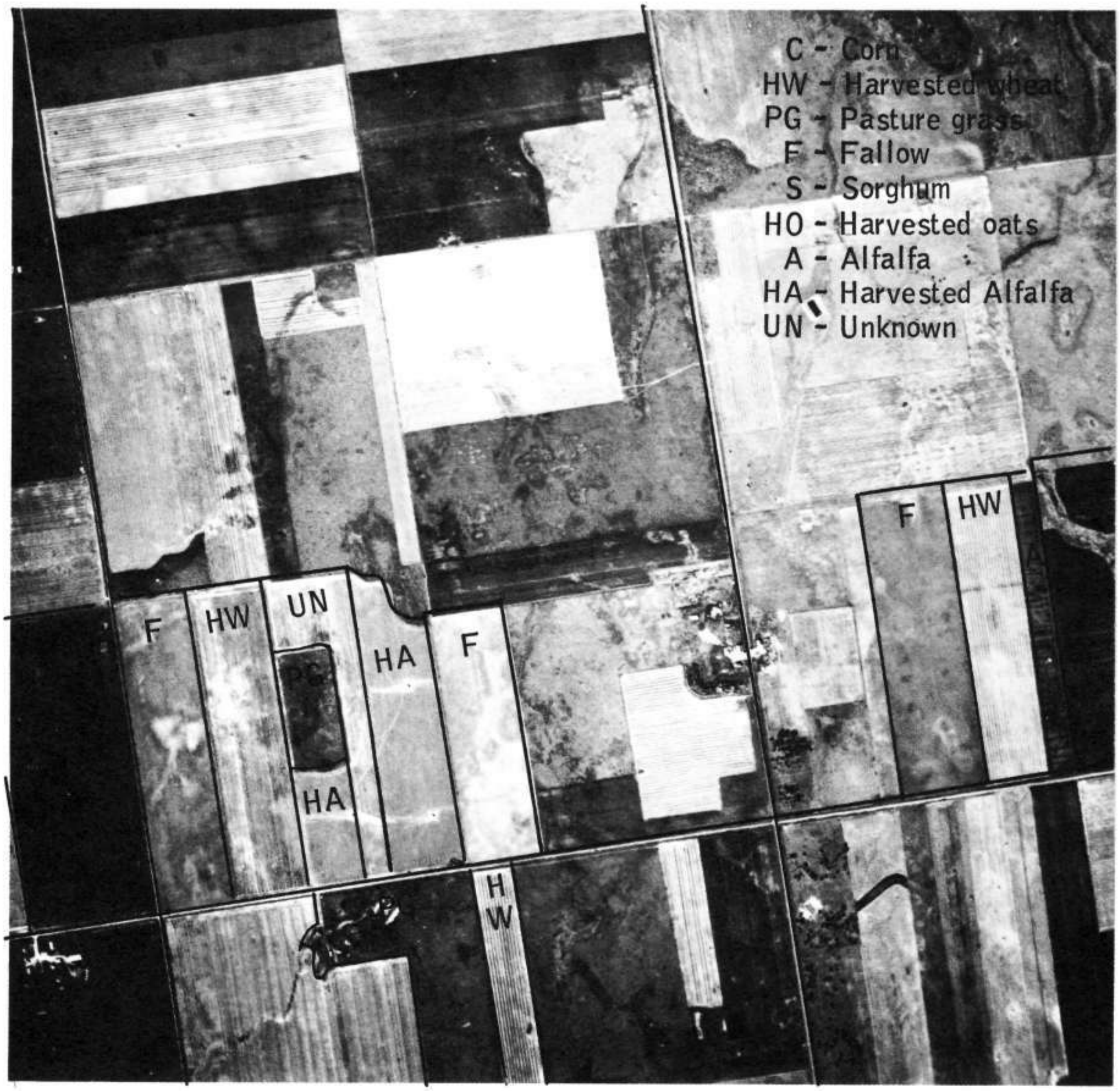


Figure 3. Remote Sensing Institute imagery taken over South Dakota at an altitude of 14,000 feet.

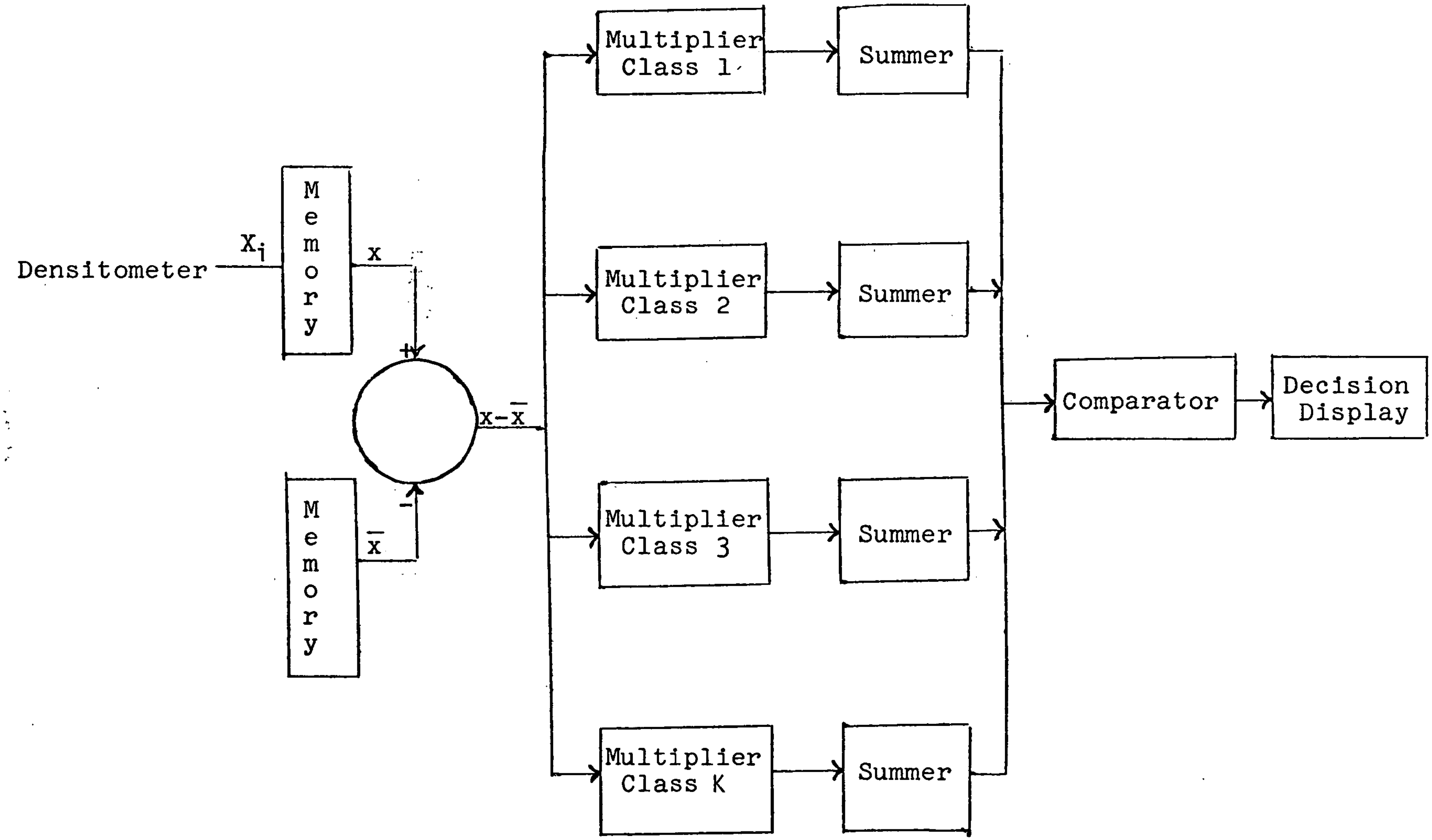


Figure 4. Special purpose K-Class classifier.

133

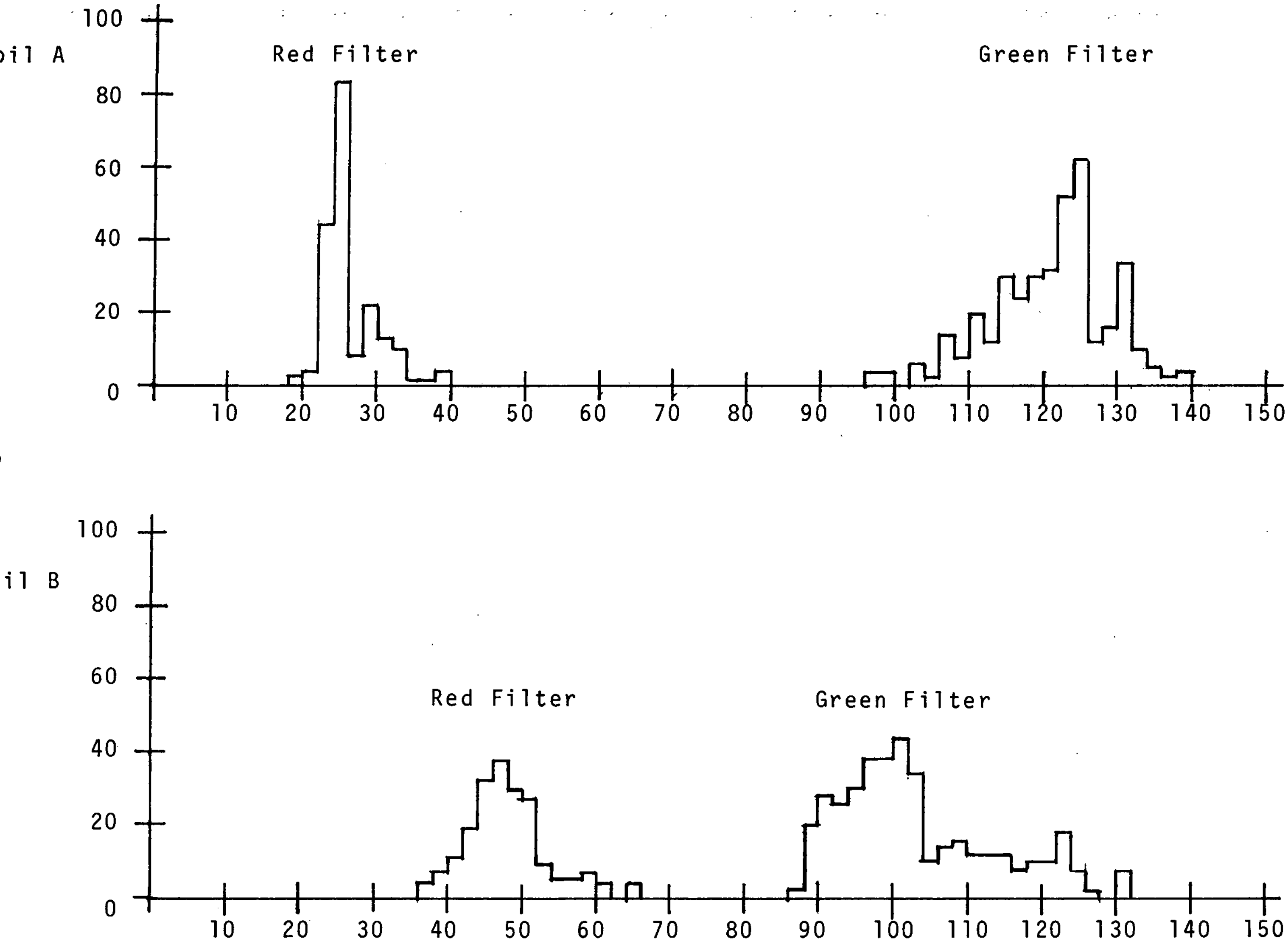


Figure 5a, Sample Probability Density Function of Soils A and B

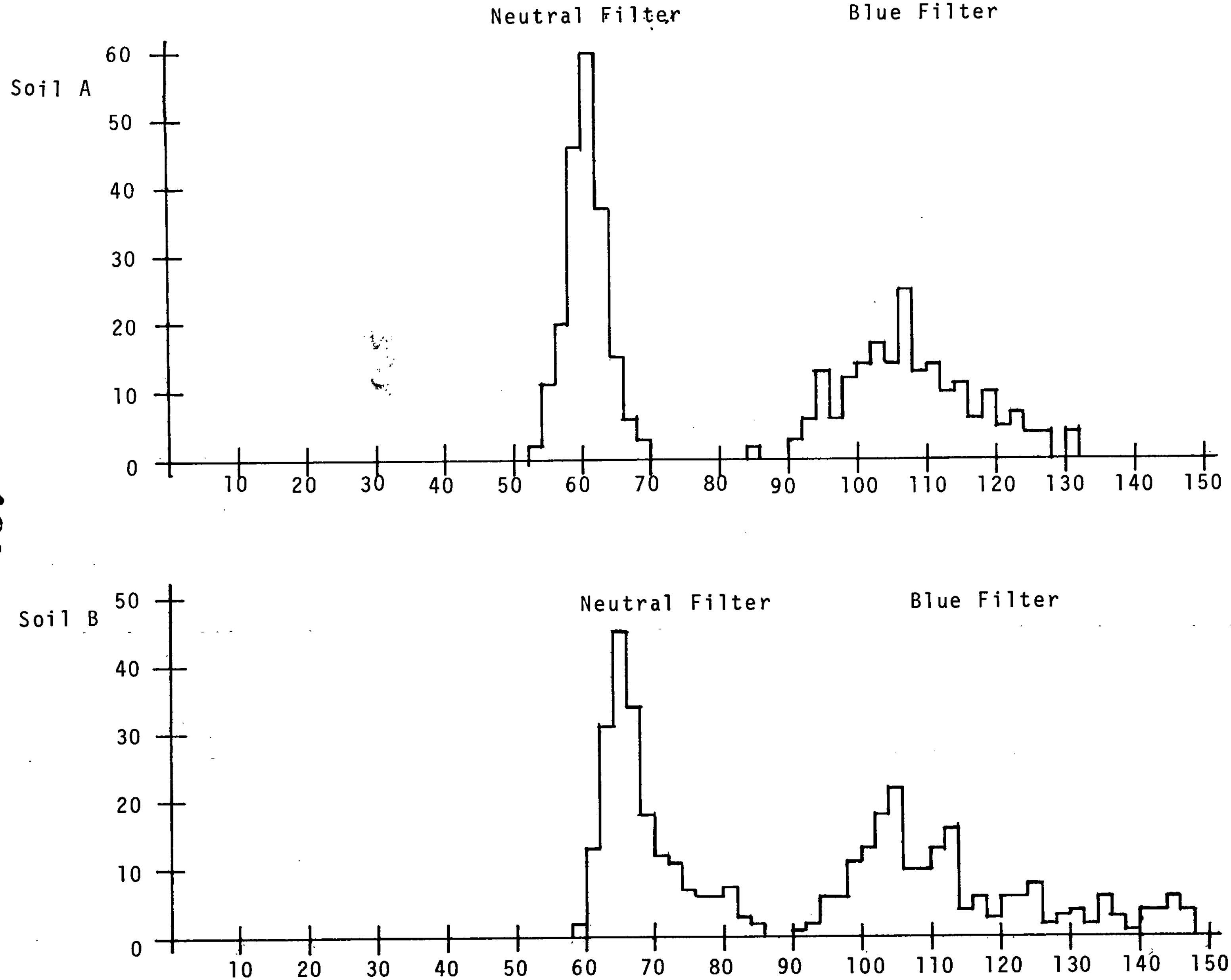
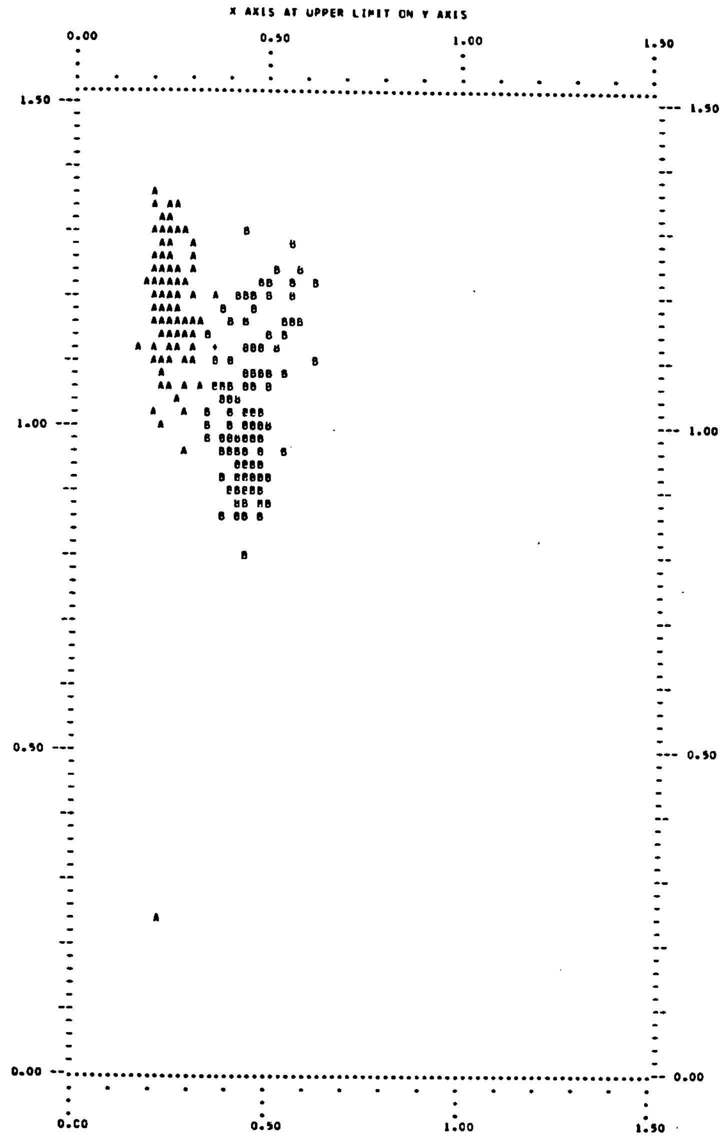
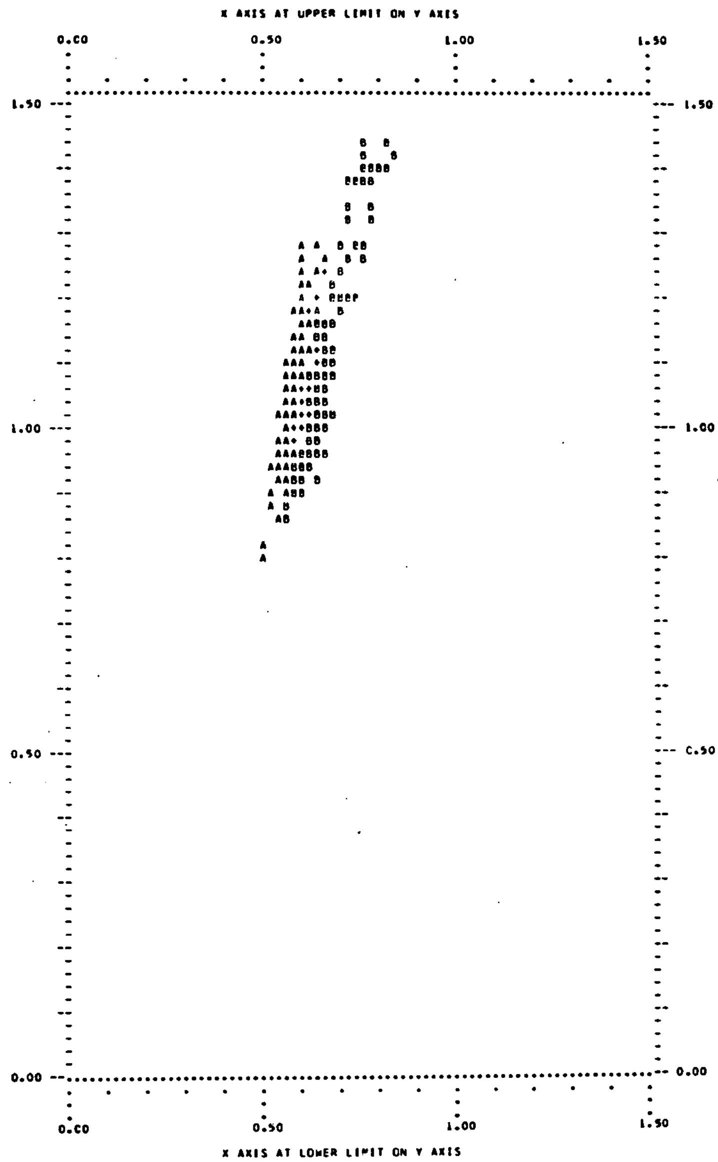


Figure 5b, Sample Probability Density Function of Soils A and B.



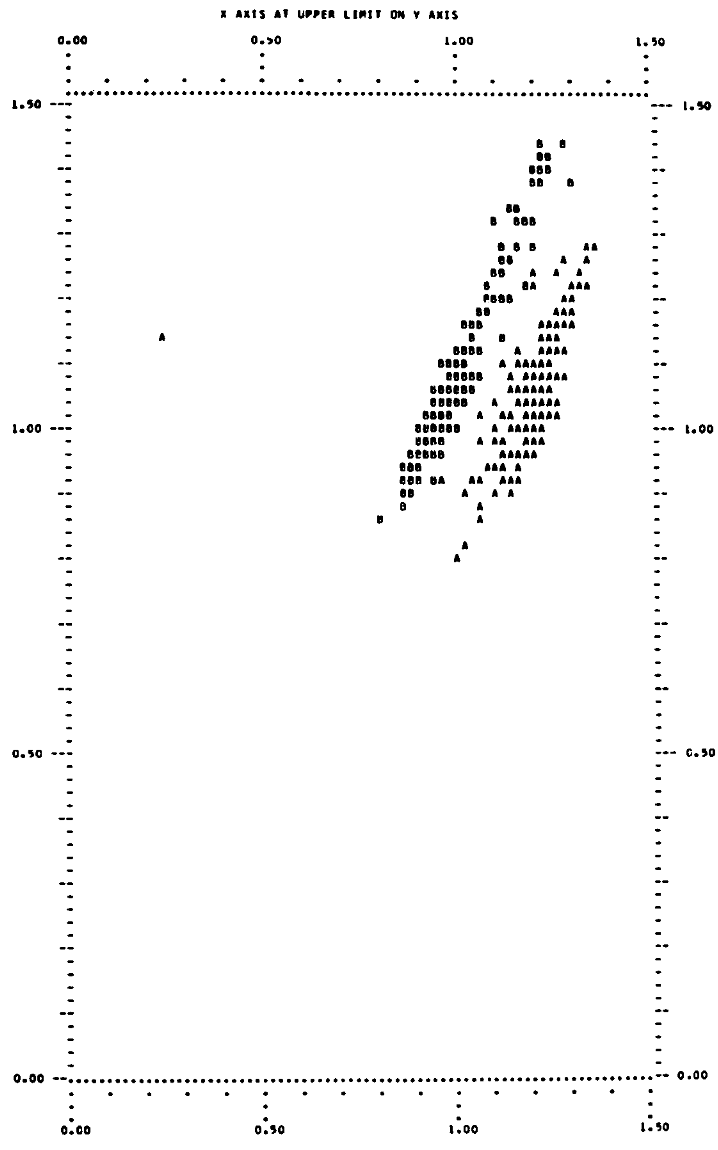
x-axis--red filter, y-axis--green filter EKIR

Figure 6. Scatter plots of density measurements for soils A and B.



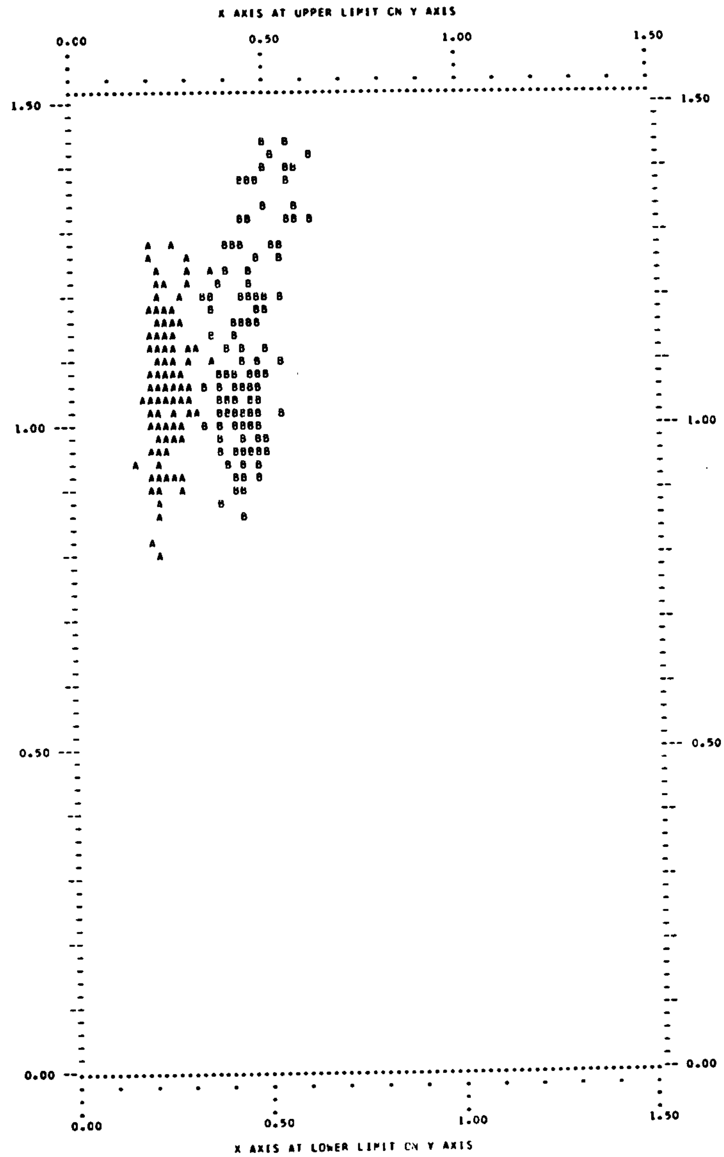
x-axis--neutral filter, y-axis--blue filter EKIR

Figure 6. Continued.



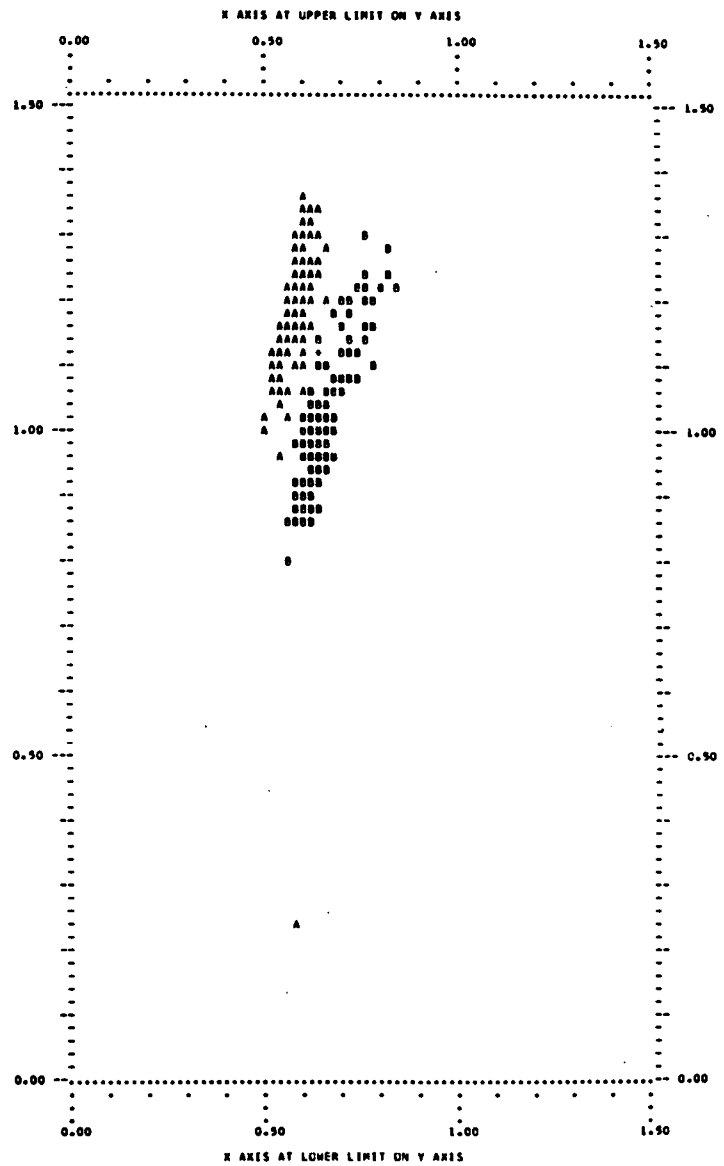
x-axis--green filter, y-axis--blue filter EKIR

Figure 6. Continued.



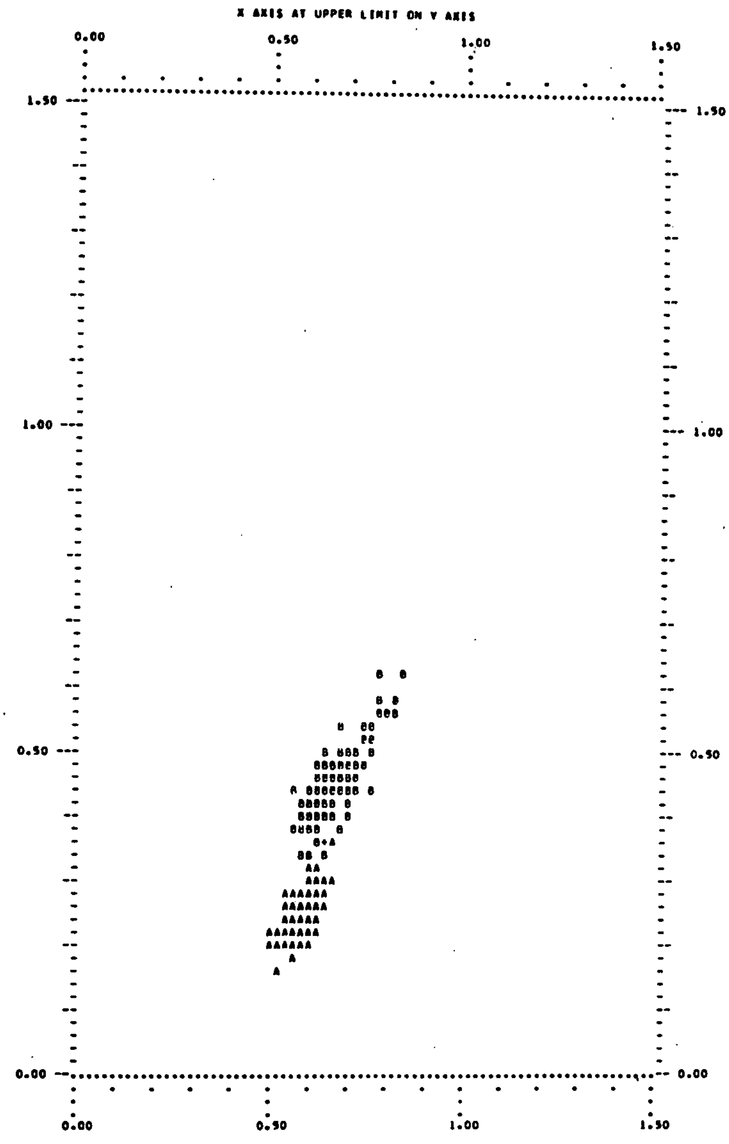
x-axis--red filter, y-axis--blue filter EKIR

Figure 6. Continued.



x-axis--neutral filter, y-axis--green filter EKIR

Figure 6. Continued.



x-axis--neutral filter, y-axis--red filter EKIR

Figure 6. Concluded.

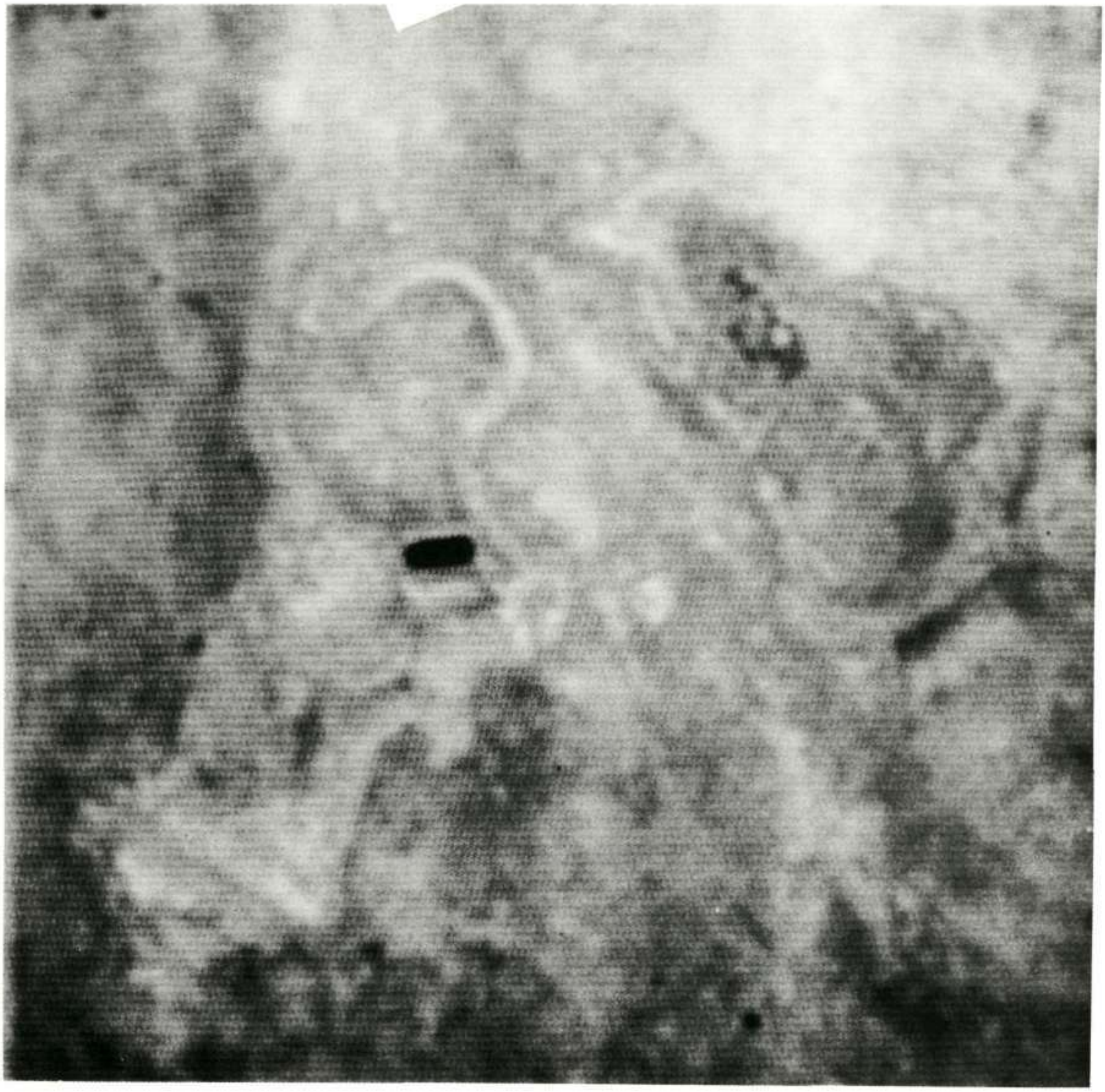


Figure 7. Spatial Data black and white image of soils A and B with neutral filter on the vidicon. The classifier was only 83 percent correct.

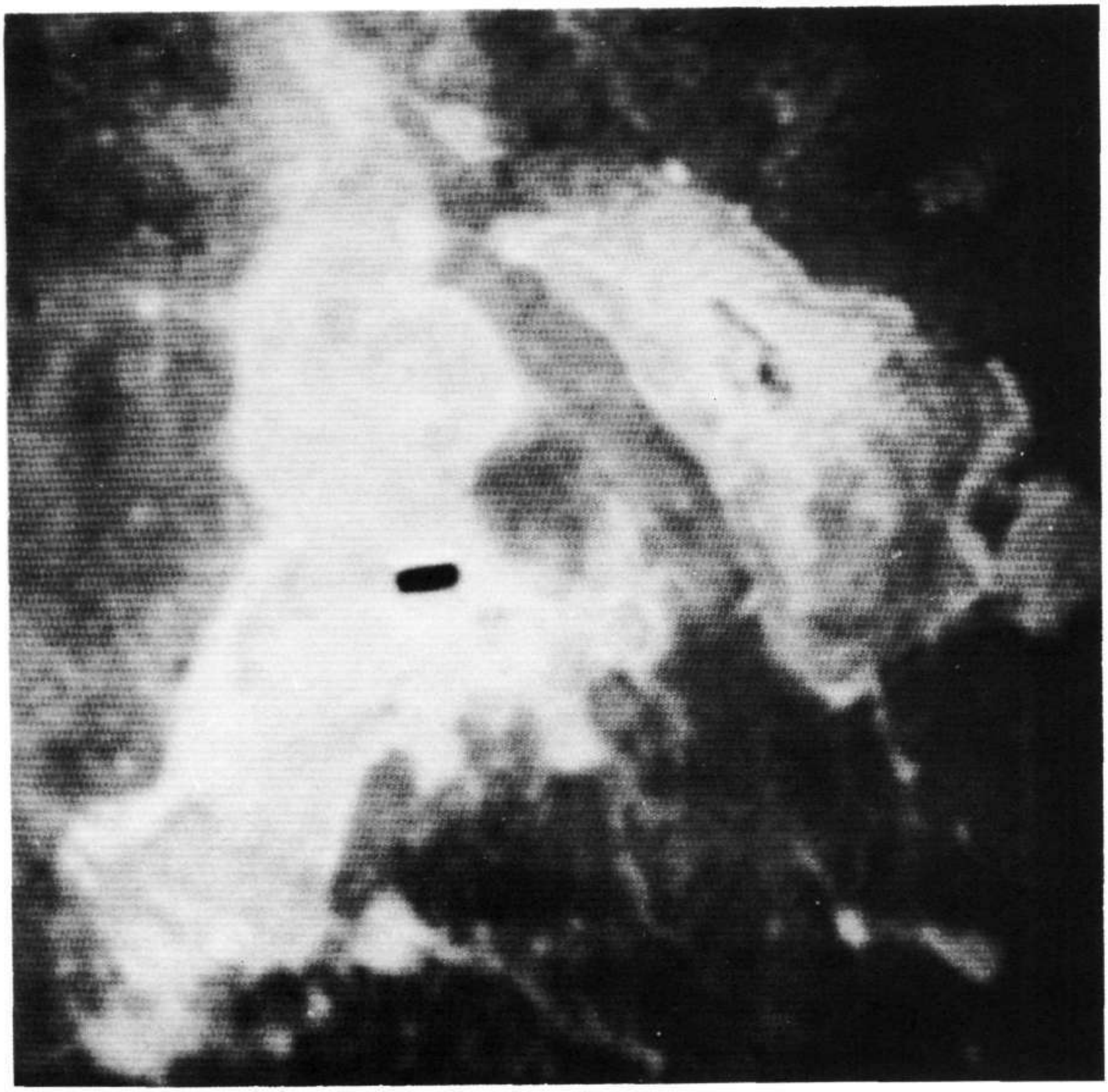


Figure 8. Spatial Data black and white image of soils A and B with a red filter on the vidicon. The classifier was 99 percent correct.

SOILS A AND B

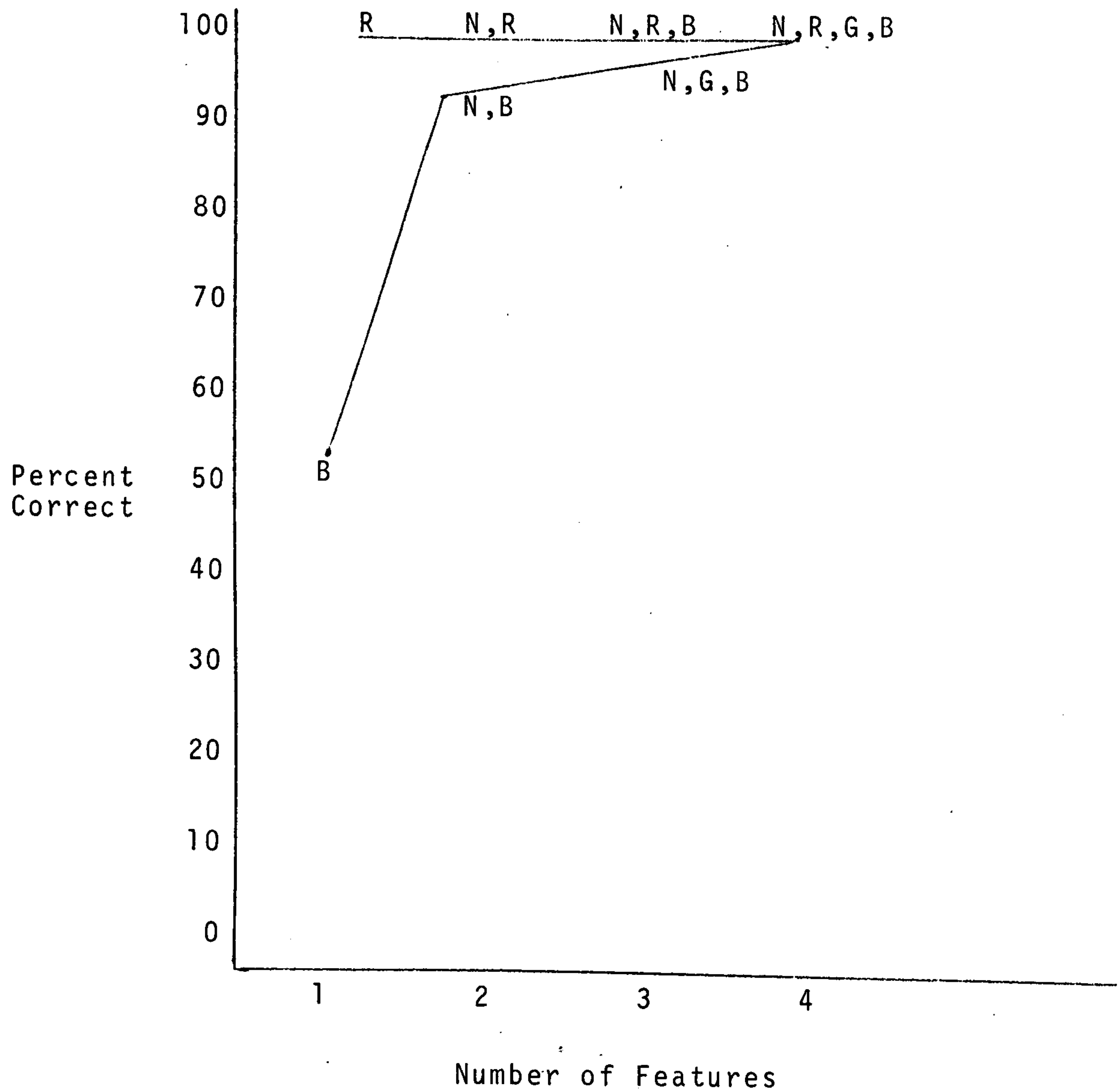


Figure 9. Classification results versus number of features.

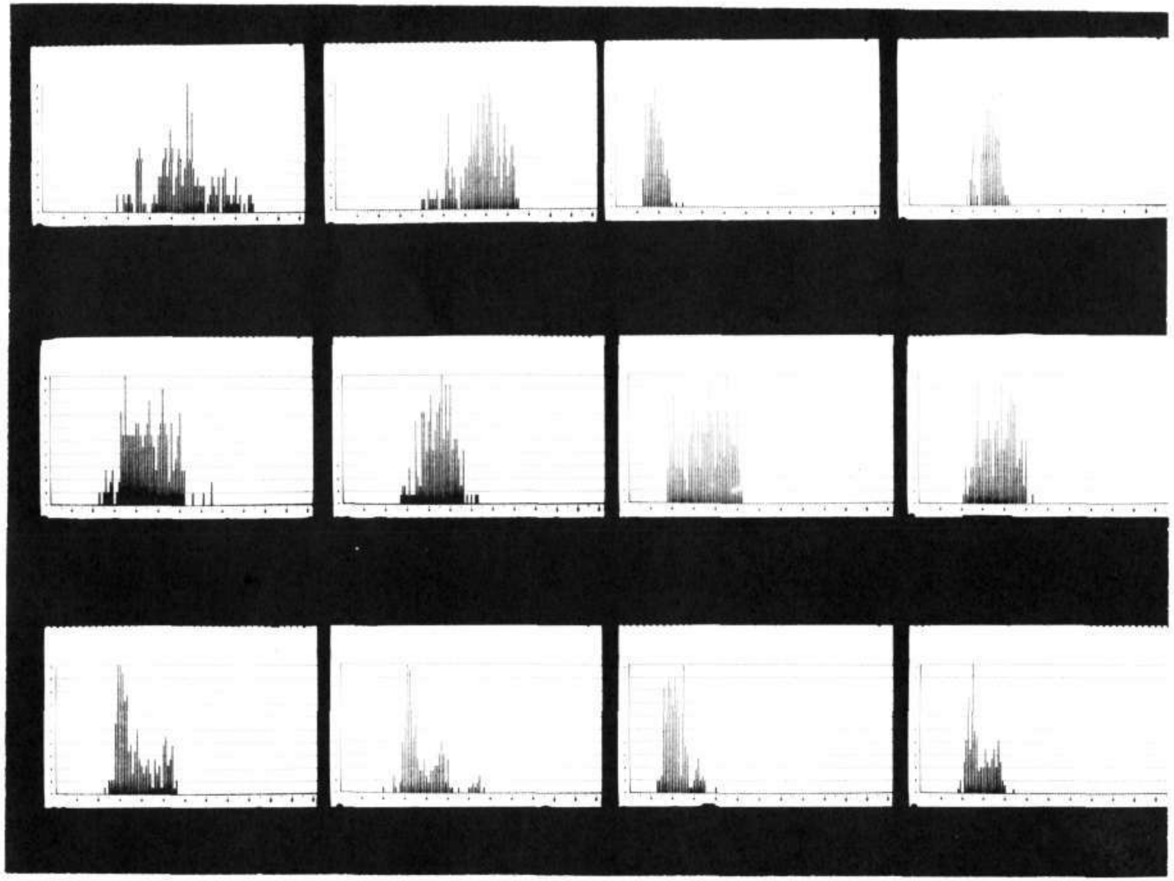
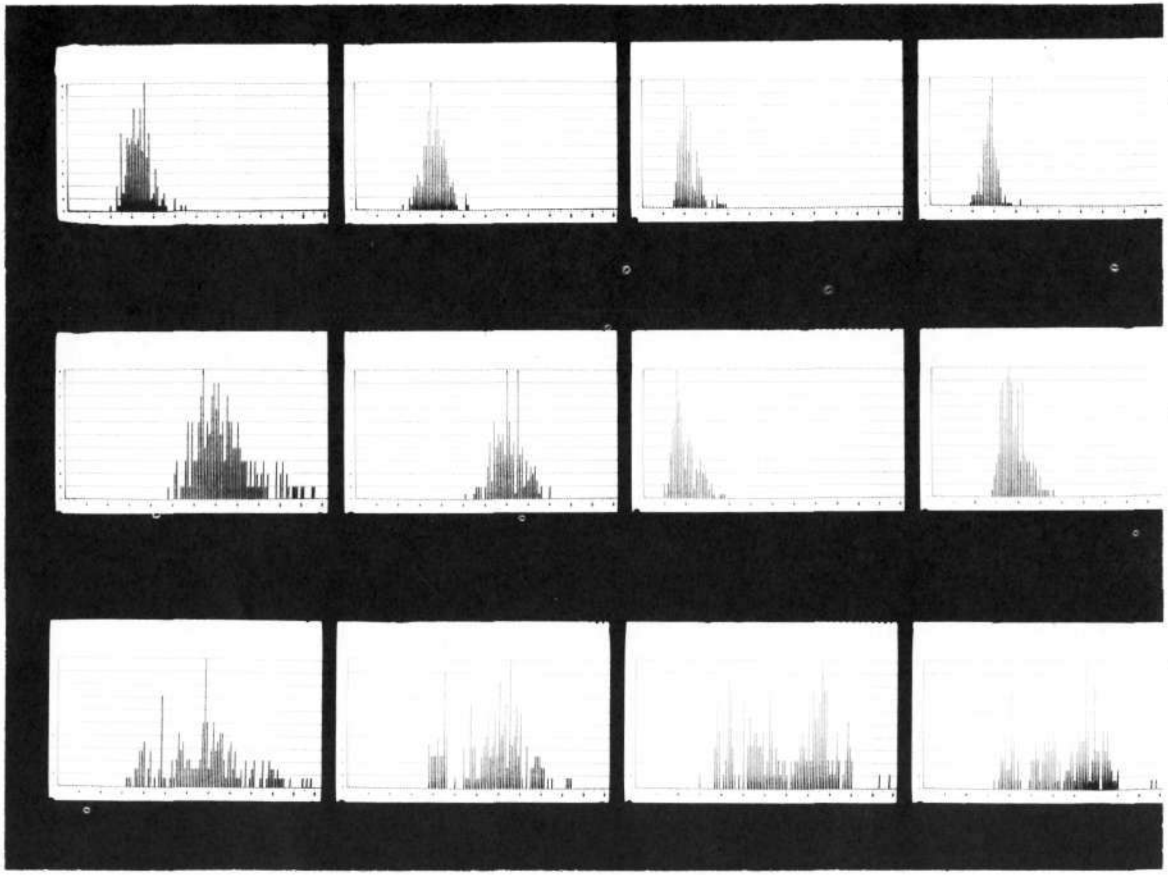


Figure 10. Sample probability density functions for the four filters and six classes. The classifier results are 88 percent correct.



Roadways

Trees

Water

Green Red Infrared Total Spectral Region

Figure 10. Concluded.